SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: Davis, Roger J.
               Raingeaud, Joel
               Gupta, Shashi
               Derijard, Benoit

    (ii) TITLE OF INVENTION: CYTOKINE-, STRESS-, AND
               ONCOPROTEIN-ACTIVATED HUMAN PROTEIN KINASE
               KINASES

    (iii) NUMBER OF SEQUENCES: 16

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: Fish & Richardson P.C.
        (B) STREET: 225 Franklin Street
        (C) CITY: Boston
        (D) STATE: MA
        (E) COUNTRY: USA
        (F) ZIP: 02110-2804

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER: US 08/530,950
        (B) FILING DATE: 19-SEP-1995
        (C) CLASSIFICATION:

    (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: Fasse, J. Peter
        (B) REGISTRATION NUMBER: 32,983
        (C) REFERENCE/DOCKET NUMBER: 04020/090001

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: 617/542-5070
        (B) TELEFAX: 617/542-8906
        (C) TELEX: 200154

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 2030 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TGGCTGGCAA TGGCCTTGCT GACCTCGAGC CGGGCCCACG TGGGGACCTT TGGAGCACAG   60

CCTACGATCC TGGTGCAAGG CCGGTGGATG CAGAGGCCAG TCCATATACC ACCCAGGCCT   120

GCGAGGAGCG TGGTCCCCAC CCATCCAGCC CATATGTGCA AGTGCCCTTG ACAGAGAGGC   180

```
TGGTCATATC CATGGTGACC ATTTATGGGC CACAACAGGT CCCCATCTGC GCAGTGAACC      240

CTGTGCTGAG CACCTTGCAG ACGTGATCTT GCTTCGTCCT GCAGCACTGT GCGGGGCAGG      300

AAAATCCAAG AGGAAGAAGG ATCTACGGAT ATCCTGCATG TCCAAGCCAC CCGCACCCAA      360

CCCCACACCC CCGCGGAACC TGGACTCCCG GACCTTCATC ACCATTGGAG ACAGAAACTT      420

TGAGGTGGAG GCTGATGACT TGGTGACCAT CTCAGAACTG GGCCGTGGAG CCTATGGGGT      480

GGTAGAGAAG GTGCGGCACG CCCAGAGCGG CACCATCATG GCCGTGAAGC GGATCCGGGC      540

CACCGTGAAC TCACAGGAGC AGAAGCGGCT GCTCATGGAC CTGGACATCA ACATGCGCAC      600

GGTCGACTGT TTCTACACTG TCACCTTCTA CGGGGCACTA TTCAGAGAGG GAGACGTGTG      660

GATCTGCATG GAGCTCATGG ACACATCCTT GGACAAGTTC TACCGGAAGG TGCTGGATAA      720

AAACATGACA ATTCCAGAGG ACATCCTTGG GGAGATTGCT GTGTCTATCG TGCGGGCCCT      780

GGAGCATCTG CACAGCAAGC TGTCGGTGAT CCACAGAGAT GTGAAGCCCT CCAATGTCCT      840

TATCAACAAG GAGGGCCATG TGAAGATGTG TGACTTTGGC ATCAGTGGCT ACTTGGTGGA      900

CTCTGTGGCC AAGACGATGG ATGCCGGCTG CAAGCCCTAC ATGGCCCCTG AGAGGATCAA      960

CCCAGAGCTG AACCAGAAGG CTACAATGT CAAGTCCGAC GTCTGGAGCC TGGGCATCAC      1020

CATGATTGAG ATGGCCATCC TGCGGTTCCC TTACGAGTCC TGGGGGACCC CGTTCCAGCA      1080

GCTGAAGCAG GTGGTGGAGG AGCCGTCCCC CCAGCTCCCA GCCGACCGTT TCTCCCCCGA      1140

GTTTGTGGAC TTCACTGCTC AGTGCCTGAG GAAGAACCCC GCAGAGCGTA TGAGCTACCT      1200

GGAGCTGATG GAGCACCCCT TCTTCACCTT GCACAAAACC AAGAAGACGG ACATTGCTGC      1260

CTTCGTGAAG AAGATCCTGG GAGAAGACTC ATAGGGGCTG GGCCTCGGAC CCCACTCCGG      1320

CCCTCCAGAG CCCCACAGCC CCATCTGCGG GGGCAGTGCT CACCCACACC ATAAGCTACT      1380

GCCATCCTGG CCCAGGGCAT CTGGGAGGAA CCGAGGGGGC TGCTCCCACC TGGCTCTGTG      1440

GCGAGCCATT TGTCCCAAGT GCCAAAGAAG CAGACCATTG GGGCTCCCAG CCAGGCCCTT      1500

GTCGGCCCCA CCAGTGCCTC TCCCTGCTGC TCCTAGGACC CGTCTCCAGC TGCTGAGATC      1560

CTGGACTGAG GGGGCCTGGA TGCCCCCTGT GGATGCTGCT GCCCCTGCAC AGCAGGCTGC      1620

CAGTGCCTGG GTGGATGGGC CACCGCCTTG CCCAGCCTGG ATGCCATCCA AGTTGTATAT      1680

TTTTTTAATC TCTCGACTGA ATGGACTTTG CACACTTTGG CCCAGGGTGG CCACACCTCT      1740

ATCCCGGCTT TGGTGCGGGG TACACAAGAG GGGATGAGTT GTGTGAATAC CCCAAGACTC      1800

CCATGAGGGA GATGCCATGA GCCGCCCAAG GCCTTCCCCT GGCACTGGCA AACAGGGCCT      1860

CTGCGGAGCA CACTGGCTCA CCCAGTCCTG CCCGCCACCG TTATCGGTGT CATTCACCTT      1920

TCGTGTTTTT TTTAATTTAT CCTCTGTTGA TTTTTTCTTT TGCTTTATGG GTTTGGCTTG      1980

TTTTTCTTGC ATGGTTTGGA GCTGATCGCT TCTCCCCCAC CCCCTAGGGG               2030
```

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 318 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: not relevant
        (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Ser Lys Pro Pro Ala Pro Asn Pro Thr Pro Pro Arg Asn Leu Asp
1               5                   10                  15

Ser Arg Thr Phe Ile Thr Ile Gly Asp Arg Met Phe Glu Val Glu Ala
            20                  25                  30

Asp Asp Leu Val Thr Ile Ser Glu Leu Gly Arg Gly Ala Tyr Gly Val
        35                  40                  45

Val Glu Lys Val Arg His Ala Gln Ser Gly Thr Ile Met Ala Val Lys
    50                  55                  60

Arg Ile Arg Ala Thr Val Asn Ser Gln Glu Gln Lys Arg Leu Leu Met
65                  70                  75                  80

Asp Leu Asp Ile Asn Met Arg Thr Val Asp Cys Phe Tyr Thr Val Thr
                85                  90                  95

Phe Tyr Gly Ala Leu Phe Arg Glu Gly Asp Val Trp Ile Cys Met Glu
            100                 105                 110

Leu Met Asp Thr Ser Leu Asp Lys Phe Tyr Arg Lys Val Leu Asp Lys
        115                 120                 125

Asn Met Thr Ile Pro Glu Asp Ile Leu Gly Glu Ile Ala Val Ser Ile
    130                 135                 140

Val Arg Ala Leu Glu His Leu His Ser Lys Leu Ser Val Ile His Arg
145                 150                 155                 160

Asp Val Lys Pro Ser Asn Val Leu Ile Asn Lys Glu Gly His Val Lys
            165                 170                 175

Met Cys Asp Phe Gly Ile Ser Gly Tyr Leu Val Asp Ser Val Ala Lys
            180                 185                 190

Thr Met Asp Ala Gly Cys Lys Pro Tyr Met Ala Pro Glu Arg Ile Asn
        195                 200                 205

Pro Glu Leu Asn Gln Lys Gly Tyr Asn Val Lys Ser Asp Val Trp Ser
    210                 215                 220

Leu Gly Ile Thr Met Ile Glu Met Ala Ile Leu Arg Phe Pro Tyr Glu
225                 230                 235                 240

Ser Trp Gly Thr Pro Phe Gln Gln Leu Lys Gln Val Val Glu Glu Pro
            245                 250                 255

Ser Pro Gln Leu Pro Ala Asp Arg Phe Ser Pro Glu Phe Val Asp Phe
        260                 265                 270

Thr Ala Gln Cys Leu Arg Lys Asn Pro Ala Glu Arg Met Ser Tyr Leu
```

```
Glu Leu Met Glu His Pro Phe Phe Thr Leu His Lys Thr Lys Lys Thr
    290                 295                 300

Asp Ile Ala Ala Phe Val Lys Lys Ile Leu Gly Glu Asp Ser
305          .          310             315
```

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1602 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
TAGCTGCAGC ACAGCCTTCC CTAACGTTGC AACTGGGGGA AAAATCACTT TCCAGTCTGT    60

TTTGCAAGGT GTGCATTTCC ATCTTGATTC CCTGAAAGTC CATCTGCTGC ATCGGTCAAG   120

AGAAACTCCA CTTGCATGAA GATTGCACGC CTGCAGCTTG CATCTTTGTT GCAAAACTAG   180

CTACAGAAGA GAAGCAAGGC AAAGTCTTTT GTGCTCCCCT CCCCCATCAA AGGAAAGGGG   240

AAAATGTCTC AGTCGAAAGG CAAGAAGCGA AACCCTGGCC TTAAAATTCC AAAAGAAGCA   300

TTTGAACAAC CTCAGACCAG TTCCACACCA CCTAGAGATT TAGACTCCAA GGCTTGCATT   360

TCTATTGGAA ATCAGAACTT TGAGGTGAAG GCAGATGACC TGGAGCCTAT AATGGAACTG   420

GGACGAGGTG CGTACGGGGT GGTGGAGAAG ATGCGGCACG TGCCCAGCGG GCAGATCATG   480

GCAGTGAAGC GGATCCGAGC CACAGTAAAT AGCCAGGAAC AGAAACGGCT ACTGATGGAT   540

TTGGATATTT CCATGAGGAC GGTGGACTGT CCATTCACTG TCACCTTTTA TGGCGCACTG   600

TTTCGGGAGG GTGATGTGTG GATCTGCATG GAGCTCATGG ATACATCACT AGATAAATTC   660

TACAAACAAG TTATTGATAA AGGCCAGACA ATTCCAGAGG ACATCTTAGG GAAAATAGCA   720

GTTTCTATTG TAAAAGCATT AGAACATTTA CATAGTAAGC TGTCTGTCAT TCACAGAGAC   780

GTCAAGCCTT CTAATGTACT CATCAATGCT CTCGGTCAAG TGAAGATGTG CGATTTTGGA   840

ATCAGTGGCT ACTTGGTGGA CTCTGTTGCT AAAACAATTG ATGCAGGTTG CAAACCATAC   900

ATGGCCCCTG AAAGAATAAA CCCAGAGCTC AACCAGAAGG GATACAGTGT GAAGTCTGAC   960

ATTTGGAGTC TGGGCATCAC GATGATTGAG TTGGCCATCC TTCGATTTCC CTATGATTCA  1020

TGGGGAACTC CATTTCAGCA GCTCAAACAG GTGGTAGAGG AGCCATCGCC ACAACTCCCA  1080

GCAGACAAGT TCTCTGCAGA GTTTGTTGAC TTTACCTCAC AGTGCTTAAA GAAGAATTCC  1140

AAAGAACGGC CTACATACCC AGAGCTAATG CAACATCCAT TTTTCACCCT ACATGAATCC  1200

AAAGGAACAG ATGTGGCATC TTTTGTAAAA CTGATTCTTG GAGACTAAAA AGCAGTGGAC  1260

TTAATCGGTT GACCCTACTG TGGATTGGTG GGTTTCGGGG TGAAGCAAGT TCACTACAGC  1320
```

ATCAATAGAA AGTCATCTTT GAGATAATTT AACCCTGCCT CTCAGAGGGT TTTCTCTCCC   1380

AATTTTCTTT TTACTCCCCC TCTTAAGGGG GCCTTGGAAT CTATAGTATA GAATGAACTG   1440

TCTAGATGGA TGAATTATGA TAAAGGCTTA GGACTTCAAA AGGTGATTAA ATATTTAATG   1500

ATGTGTCATA TGAGTCCTCA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA   1560

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA   1602

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 334 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: not relevant
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Ser Gln Ser Lys Gly Lys Lys Arg Asn Pro Gly Leu Lys Ile Pro
1               5                   10                  15

Lys Glu Ala Phe Glu Gln Pro Gln Thr Ser Ser Thr Pro Pro Arg Asp
            20                  25                  30

Leu Asp Ser Lys Ala Cys Ile Ser Ile Gly Asn Gln Asn Phe Glu Val
        35                  40                  45

Lys Ala Asp Asp Leu Glu Pro Ile Met Glu Leu Gly Arg Gly Ala Tyr
50                  55                  60

Gly Val Val Glu Lys Met Arg His Val Pro Ser Gly Gln Ile Met Ala
65                  70                  75                  80

Val Lys Arg Ile Arg Ala Thr Val Asn Ser Gln Glu Gln Lys Arg Leu
            85                  90                  95

Leu Met Asp Leu Asp Ile Ser Met Arg Thr Val Asp Cys Pro Phe Thr
            100                 105                 110

Val Thr Phe Tyr Gly Ala Leu Phe Arg Glu Gly Asp Val Trp Ile Cys
            115                 120                 125

Met Glu Leu Met Asp Thr Ser Leu Asp Lys Phe Tyr Lys Gln Val Ile
    130                 135                 140

Asp Lys Gly Gln Thr Ile Pro Glu Asp Ile Leu Gly Lys Ile Ala Val
145                 150                 155                 160

Ser Ile Val Lys Ala Leu Glu His Leu His Ser Lys Leu Ser Val Ile
                165                 170                 175

His Arg Asp Val Lys Pro Ser Asn Val Leu Ile Asn Ala Leu Gly Gln
                180                 185                 190

Val Lys Met Cys Asp Phe Gly Ile Ser Gly Tyr Leu Val Asp Ser Val
            195                 200                 205

Ala Lys Thr Ile Asp Ala Gly Cys Lys Pro Tyr Met Ala Pro Glu Arg
            210                 215                 220
```

- 55 -

```
Ile Asn Pro Glu Leu Asn Gln Lys Gly Tyr Ser Val Lys Ser Asp Ile
225                 230             235                 240

Trp Ser Leu Gly Ile Thr Met Ile Glu Leu Ala Ile Leu Arg Phe Pro
                245             250                 255

Tyr Asp Ser Trp Gly Thr Pro Phe Gln Gln Leu Lys Gln Val Val Glu
            260             265                 270

Glu Pro Ser Pro Gln Leu Pro Ala Asp Lys Phe Ser Ala Glu Phe Val
        275             280                 285

Asp Phe Thr Ser Gln Cys Leu Lys Lys Asn Ser Lys Glu Arg Pro Thr
        290             295                 300

Tyr Pro Glu Leu Met Gln His Pro Phe Phe Thr Leu His Glu Ser Lys
305             310                 315                 320

Gly Thr Asp Val Ala Ser Phe Val Lys Leu Ile Leu Gly Asp
                325             330
```

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 3497 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
CTAGGGTCCC CGGCGCCAGG CCACCCGGCC GTCAGCAGCA TGCAGGGTAA ACGCAAAGCA     60

CTGAAGTTGA ATTTTGCAAA TCCACCTTTC AAATCTACAG CAAGGTTTAC TCTGAATCCC    120

AATCCTACAG GAGTTCAAAA CCCACACATA GAGAGACTGA GAACACACAG CATTGAGTCA    180

TCAGGAAAAC TGAAGATCTC CCCTGAACAA CACTGGGATT TCACTGCAGA GGACTTGAAA    240

GACCTTGGAG AAAATTGGACG AGGAGCTTAT GGTTCTGTCA ACAAAATGGT CCACAAACCA    300

AGTGGGCAAA TAATGGCAGT TAAAAGAATT CGGTCAACAG TGGATGAAAA AGAACAAAAA    360

CAACTTCTTA TGGATTTGGA TGTAGTAATG CGGAGTAGTG ATTGCCCATA CATTGTTCAG    420

TTTTATGGTG CACTCTTCAG AGAGGGTGAC TGTTGGATCT GTATGGAACT CATGTCTACC    480

TCGTTTGATA AGTTTTACAA ATATGTATAT AGTGTATTAG ATGATGTTAT TCCAGAAGAA    540

ATTTTAGGCA AAATCACTTT AGCAACTGTG AAAGCACTAA ACCACTTAAA AGAAAACTTG    600

AAAATTATTC ACAGAGATAT CAAACCTTCC AATATTCTTC TGGACAGAAG TGGAAATATT    660

AAGCTCTGTG ACTTCGGCAT CAGTGGACAG CTTGTGGACT CTATTGCCAA GACAAGAGAT    720

GCTGGCTGTA GGCCATACAT GGCACCTGAA AGAATAGACC CAAGCGCATC ACGACAAGGA    780

TATGATGTCC GCTCTGATGT CTGGAGTTTG GGGATCACAT TGTATGAGTT GGCCACAGGC    840

CGATTTCCTT ATCCAAAGTG GAATAGTGTA TTTGATCAAC TAACACAAGT CGTGAAAGGA    900

GATCCTCCGC AGCTGAGTAA TTCTGAGGAA AGGGAATTCT CCCCGAGTTT CATCAACTTT    960
```

- 56 -

```
GTCAACTTGT GCCTTACGAA GGATGAATCC AAAAGGCCAA AGTATAAAGA GCTTCTGAAA    1020

CATCCCTTTA TTTTGATGTA TGAAGAACGT GCCGTTGAGG TCGCATGCTA TGTTTGTAAA    1080

ATCCTGGATC AAATGCCAGC TACTCCCAGC TCTCCCATGT ATGTCGATTG ATATCGTGCT    1140

ACATCAGACT CTAGAAAAAA GGGCTGAGAG GAAGCAAGAC GTAAAGAATT TTCATCCCGT    1200

ATCACAGTGT TTTTATTGCT CGCCCAGACA CCATGTGCAA TAAGATTGGT GTTCGTTTCC    1260

ATCATGTCTG TATACTCCTG TCACCTAGAA CGTGCATCCT TGTAATACCT GATTGATCAC    1320

ACAGTGTTAG TGCTGGTCAG AGAGACCTCA TCCTGCTCTT TTGTGATGAA CATATTCATG    1380

AAATGTGGAA GTCAGTACGA TCAAGTTGTT GACTGTGATT AGATCACATC TTAAATTCAT    1440

TTCTAGACTC AAAACCTGGA GATGCAGCTA CTGGAATGGT GTTTTGTCAG ACTTCCAAAT    1500

CCTGGAAGGA CACAGTGATG AATGTACTAT ATCTGAACAT AGAAACTCGG GCTTGAGTGA    1560

GAAGAGCTTG CACAGCCAAC GAGACACATT GCCTTCTGGA GCTGGGAGAC AAAGGAGGAA    1620

TTTACTTTCT TCACCAAGTG CAATAGATTA CTGATGTGAT ATTCTGTTGC TTTACAGTTA    1680

CAGTTGATGT TTGGGGATCG ATGTGCTCAG CCAAATTTCC TGTTTGAAAT ATCATGTTAA    1740

ATTAGAATGA ATTTATCTTT ACCAAAAACC ATGTTGCGTT CAAAGAGGTG AACATTAAAA    1800

TATAGAGACA GGACAGAATG TGTTCTTTTC TCCTCTACCA GTCCTATTTT TCAATGGGAA    1860

GACTCAGGAG TCTGCCACTT GTCAAAGAAG GTGCTGATCC TAAGAATTTT TCATTCTCAG    1920

AATTCGGTGT GCTGCCAACT TGATGTTCCA CCTGCCACAA ACCACCAGGA CTGAAAGAAG    1980

AAAACAGTAC AGAAGGCAAA GTTTACAGAT GTTTTTAATT CTAGTATTTT ATCTGGAACA    2040

ACTTGTAGCA GCTATATATT TCCCCTTGGT CCCAAGCCTG ATACTTTAGC CATCATAACT    2100

CACTAACAGG GAGAAGTAGC TAGTAGCAAT GTGCCTTGAT TGATTAGATA AAGATTTCTA    2160

GTAGGCAGCA AAAGACCAAA TCTCAGTTGT TTGCTTCTTG CCATCACTGG TCCAGGTCTT    2220

CAGTTTCCGA ATCTCTTTCC CTTCCCCTGT GGTCTATTGT CGCTATGTGA CTTGCGCTTA    2280

ATCCAATATT TTGCCTTTTT TCTATATCAA AAAACCTTTA CAGTTAGCAG GGATGTTCCT    2340

TACCGAGGAT TTTTAACCCC CAATCTCTCA TAATCGCTAG TGTTTAAAAG GCTAAGAATA    2400

GTGGGGCCCA ACCGATGTGG TAGGTGATAA AGAGGCATCT TTTCTAGAGA CACATTGGAC    2460

CAGATGAGGA TCCGAAACGG CAGCCTTTAC GTTCATCACC TGCTAGAACC TCTCGTAGTC    2520

CATCACCATT TCTTGGCATT GGAATTCTAC TGGAAAAAAA TACAAAAAGC AAAACAAAAC    2580

CCTCAGCACT GTTACAAGAG GCCATTTAAG TATCTTGTGC TTCTTCACTT ACCCATTAGC    2640

CAGGTTCTCA TTAGGTTTTG CTTGGGCCTC CCTGGCACTG AACCTTAGGC TTTGTATGAC    2700

AGTGAAGCAG CACTGTGAGT GGTTCAAGCA CACTGGAATA TAAAACAGTC ATGGCCTGAG    2760

ATGCAGGTGA TGCCATTACA GAACCAAATC GTGGCACGTA TTGCTGTGTC TCCTCTCAGA    2820
```

- 57 -

```
GTGACAGTCA TAAATACTGT CAAACAATAA AGGGAGAATG GTGCTGTTTA AAGTCACATC    2880

CCTGTAAATT GCAGAATTCA AAAGTGATTA TCTCTTTGAT CTACTTGCCT CATTTCCCTA    2940

TCTTCTCCCC CACGGTATCC TAAACTTTAG ACTTCCCACT GTTCTGAAAG GAGACATTGC    3000

TCTATGTCTG CCTTCGACCA CAGCAAGCCA TCATCCTCCA TTGCTCCCGG GGACTCAAGA    3060

GGAATCTGTT TCTCTGCTGT CAACTTCCCA TCTGGCTCAG CATAGGGTCA CTTTGCCATT    3120

ATGCAAATGG AGATAAAAGC AATTCTGGCT GTCCAGGAGC TAATCTGACC GTTCTATTGT    3180

GTGGATGACC ACATAAGAAG GCAATTTTAG TGTATTAATC ATAGATTATT ATAAACTATA    3240

AACTTAAGGG CAAGGAGTTT ATTACAATGT ATCTTTATTA AAACAAAAGG GTGTATAGTG    3300

TTCACAAACT GTGAAAATAG TGTAAGAACT GTACATTGTG AGCTCTGGTT ATTTTTCTCT    3360

TGTACCATAG AAAAATGTAT AAAAATTATC AAAAAGCTAA TGTGCAGGGA TATTGCCTTA    3420

TTTGTCTGTA AAAAATGGAG CTCAGTAACA TAACTGCTTC TTGGAGCTTT GGAATATTTT    3480

ATCCTGTATT CTTGTTT                                                   3497
```

(2) INFORMATION FOR SEQ ID NO:6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 363 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: not relevant
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Gln Gly Lys Arg Lys Ala Leu Lys Leu Asn Phe Ala Asn Pro Pro
1               5                   10                  15

Phe Lys Ser Thr Ala Arg Phe Thr Leu Asn Pro Asn Pro Thr Gly Val
        20                  25                  30

Gln Asn Pro His Ile Glu Arg Leu Arg Thr His Ser Ile Glu Ser Ser
        35                  40                  45

Gly Lys Leu Lys Ile Ser Pro Glu Gln His Trp Asp Phe Thr Ala Glu
    50                  55                  60

Asp Leu Lys Asp Leu Gly Glu Ile Gly Arg Gly Ala Tyr Gly Ser Val
65                  70                  75                  80

Asn Lys Met Val His Lys Pro Ser Gly Gln Ile Met Ala Val Lys Arg
                85                  90                  95

Ile Arg Ser Thr Val Asp Glu Lys Glu Gln Lys Gln Leu Leu Met Asp
            100                 105                 110

Leu Asp Val Val Met Arg Ser Ser Asp Cys Pro Tyr Ile Val Gln Phe
        115                 120                 125

Tyr Gly Ala Leu Phe Arg Glu Gly Asp Cys Trp Ile Cys Met Glu Leu
    130                 135                 140

Met Ser Thr Ser Phe Asp Lys Phe Tyr Lys Tyr Val Tyr Ser Val Leu
```

- 58 -

Asp Asp Val Ile Pro Glu Glu Ile Leu Gly Lys Ile Thr Leu Ala Thr
                165   170   175

Val Lys Ala Leu Asn His Leu Lys Glu Asn Leu Lys Ile Ile His Arg
            180   185   190

Asp Ile Lys Pro Ser Asn Ile Leu Leu Asp Arg Ser Gly Asn Ile Lys
        195   200   205

Leu Cys Asp Phe Gly Ile Ser Gly Gln Leu Val Asp Ser Ile Ala Lys
    210   215   220

Thr Arg Asp Ala Gly Cys Arg Pro Tyr Met Ala Pro Glu Arg Ile Asp
225   230   235   240

Pro Ser Ala Ser Arg Gln Gly Tyr Asp Val Arg Ser Asp Val Trp Ser
            245   250   255

Leu Gly Ile Thr Leu Tyr Glu Leu Ala Thr Gly Arg Phe Pro Tyr Pro
                260   265   270

Lys Trp Asn Ser Val Phe Asp Gln Leu Thr Gln Val Val Lys Gly Asp
        275   280   285

Pro Pro Gln Leu Ser Asn Ser Glu Glu Arg Glu Phe Ser Pro Ser Phe
    290   295   300

Ile Asn Phe Val Asn Leu Cys Leu Thr Lys Asp Glu Ser Lys Arg Pro
305   310   315   320

Lys Tyr Lys Glu Leu Leu Lys His Pro Phe Ile Leu Met Tyr Glu Glu
                325   330   335

Arg Ala Val Glu Val Ala Cys Tyr Val Cys Lys Ile Leu Asp Gln Met
            340   345   350

Pro Ala Thr Pro Ser Ser Pro Met Tyr Val Asp
        355   360

(2) INFORMATION FOR SEQ ID NO:7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 3553 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CAACAATGGC GGCTCCGAGC CCGAGCGGTG GCGGCGGCAG CGGCACCCCC GGCCCCGTAG   60

GGTCCCCGGC GCCAGGCCAC CCGGCCGTCA GCAGCATGCA GGGTAAACGC AAAGCACTGA   120

AGTTGAATTT TGCAAATCCA CCTTTCAAAT CTACAGCAAG GTTTACTCTG AATCCCAATC   180

CTACAGGAGT TCAAAACCCA CACATAGAGA GACTGAGAAC ACACAGCATT GAGTCATCAG   240

GAAAACTGAA GATCTCCCCT GAACAACACT GGGATTTCAC TGCAGAGGAC TTGAAAGACC   300

TTGGAGAAAT TGGACGAGGA GCTTATGGTT CTGTCAACAA AATGGTCCAC AAACCAAGTG   360

- 59 -

```
GGCAAATAAT GGCAGTTAAA AGAATTCGGT CAACAGTGGA TGAAAAAGAA CAAAAACAAC    420

TTCTTATGGA TTTGGATGTA GTAATGCGGA GTAGTGATTG CCCATACATT GTTCAGTTTT    480

ATGGTGCACT CTTCAGAGAG GGTGACTGTT GGATCTGTAT GGAACTCATG TCTACCTCGT    540

TTGATAAGTT TTACAAATAT GTATATAGTG TATTAGATGA TGTTATTCCA GAAGAAATTT    600

TAGGCAAAAT CACTTTAGCA ACTGTGAAAG CACTAAACCA CTTAAAAGAA AACTTGAAAA    660

TTATTCACAG AGATATCAAA CCTTCCAATA TTCTTCTGGA CAGAAGTGGA AATATTAAGC    720

TCTGTGACTT CGGCATCAGT GGACAGCTTG TGGACTCTAT TGCCAAGACA AGAGATGCTG    780

GCTGTAGGCC ATACATGGCA CCTGAAAGAA TAGACCCAAG CGCATCACGA CAAGGATATG    840

ATGTCCGCTC TGATGTCTGG AGTTTGGGGA TCACATTGTA TGAGTTGGCC ACAGGCCGAT    900

TTCCTTATCC AAAGTGGAAT AGTGTATTTG ATCAACTAAC ACAAGTCGTG AAAGGAGATC    960

CTCCGCAGCT GAGTAATTCT GAGGAAAGGG AATTCTCCCC GAGTTTCATC AACTTTGTCA   1020

ACTTGTGCCT TACGAAGGAT GAATCCAAAA GGCCAAAGTA TAAAGAGCTT CTGAAACATC   1080

CCTTTATTTT GATGTATGAA GAACGTGCCG TTGAGGTCGC ATGCTATGTT TGTAAAATCC   1140

TGGATCAAAT GCCAGCTACT CCCAGCTCTC CCATGTATGT CGATTGATAT CGTGCTACAT   1200

CAGACTCTAG AAAAAAGGGC TGAGAGGAAG CAAGACGTAA AGAATTTTCA TCCCGTATCA   1260

CAGTGTTTTT ATTGCTCGCC CAGACACCAT GTGCAATAAG ATTGGTGTTC GTTTCCATCA   1320

TGTCTGTATA CTCCTGTCAC CTAGAACGTG CATCCTTGTA ATACCTGATT GATCACACAG   1380

TGTTAGTGCT GGTCAGAGAG ACCTCATCCT GCTCTTTTGT GATGAACATA TTCATGAAAT   1440

GTGGAAGTCA GTACGATCAA GTTGTTGACT GTGATTAGAT CACATCTTAA ATTCATTTCT   1500

AGACTCAAAA CCTGGAGATG CAGCTACTGG AATGGTGTTT TGTCAGACTT CCAAATCCTG   1560

GAAGGACACA GTGATGAATG TACTATATCT GAACATAGAA ACTCGGGCTT GAGTGAGAAG   1620

AGCTTGCACA GCCAACGAGA CACATTGCCT TCTGGAGCTG GGAGACAAAG GAGGAATTTA   1680

CTTTCTTCAC CAAGTGCAAT AGATTACTGA TGTGATATTC TGTTGCTTTA CAGTTACAGT   1740

TGATGTTTGG GGATCGATGT GCTCAGCCAA ATTTCCTGTT TGAAATATCA TGTTAAATTA   1800

GAATGAATTT ATCTTTACCA AAAACCATGT TGCGTTCAAA GAGGTGAACA TTAAAATATA   1860

GAGACAGGAC AGAATGTGTT CTTTTCTCCT CTACCAGTCC TATTTTTCAA TGGGAAGACT   1920

CAGGAGTCTG CCACTTGTCA AAGAAGGTGC TGATCCTAAG AATTTTTCAT TCTCAGAATT   1980

CGGTGTGCTG CCAACTTGAT GTTCCACCTG CCACAAACCA CCAGGACTGA AGAAGAAAA    2040

CAGTACAGAA GGCAAAGTTT ACAGATGTTT TTAATTCTAG TATTTTATCT GGAACAACTT   2100

GTAGCAGCTA TATATTTCCC CTTGGTCCCA AGCCTGATAC TTTAGCCATC ATAACTCACT   2160

AACAGGGAGA AGTAGCTAGT AGCAATGTGC CTTGATTGAT TAGATAAAGA TTTCTAGTAG   2220
```

GCAGCAAAAG ACCAAATCTC AGTTGTTTGC TTCTTGCCAT CACTGGTCCA GGTCTTCAGT   2280

TTCCGAATCT CTTTCCCTTC CCCTGTGGTC TATTGTCGCT ATGTGACTTG CGCTTAATCC   2340

AATATTTTGC CTTTTTTCTA TATCAAAAAA CCTTTACAGT TAGCAGGGAT GTTCCTTACC   2400

GAGGATTTTT AAGCCCCAAT CTCTCATAAT CGCTAGTGTT TAAAAGGCTA AGAATAGTGG   2460

GGCCCAACCG ATGTGGTAGG TGATAAAGAG GCATCTTTTC TAGAGACACA TTGGACCAGA   2520

TGAGGATCCG AAACGGCAGC CTTTACGTTC ATCACCTGCT AGAACCTCTC GTAGTCCATC   2580

ACCATTTCTT GGCATTGGAA TTCTACTGGA AAAAAATACA AAAAGCAAAA CAAAACCCTC   2640

AGCACTGTTA CAAGAGGCCA TTTAAGTATC TTGTGCTTCT TCACTTACCC ATTAGCCAGG   2700

TTCTCATTAG GTTTTGCTTG GGCCTCCCTG GCACTGAACC TTAGGCTTTG TATGACAGTG   2760

AAGCAGCACT GTGAGTGGTT CAAGCACACT GGAATATAAA ACAGTCATGG CCTGAGATGC   2820

AGGTGATGCC ATTACAGAAC CAAATCGTGG CACGTATTGC TGTGTCTCCT CTCAGAGTGA   2880

CAGTCATAAA TACTGTCAAA CAATAAAGGG AGAATGGTGC TGTTTAAAGT CACATCCCTG   2940

TAAATTGCAG AATTCAAAAG TGATTATCTC TTTGATCTAC TTGCCTCATT TCCCTATCTT   3000

CTCCCCCACG GTATCCTAAA CTTTAGACTT CCCACTGTTC TGAAAGGAGA CATTGCTCTA   3060

TGTCTGCCTT CGACCACAGC AAGCCATCAT CCTCCATTGC TCCCGGGGAC TCAAGAGGAA   3120

TCTGTTTCTC TGCTGTCAAC TTCCCATCTG GCTCAGCATA GGGTCACTTT GCCATTATGC   3180

AAATGGAGAT AAAAGCAATT CTGGCTGTCC AGGAGCTAAT CTGACCGTTC TATTGTGTGG   3240

ATGACCACAT AAGAAGGCAA TTTTAGTGTA TTAATCATAG ATTATTATAA ACTATAAACT   3300

TAAGGGCAAG GAGTTTATTA CAATGTATCT TTATTAAAAC AAAAGGGTGT ATAGTGTTCA   3360

CAAACTGTGA AAATAGTGTA AGAACTGTAC ATTGTGAGCT CTGGTTATTT TTCTCTTGTA   3420

CCATAGAAAA ATGTATAAAA ATTATCAAAA AGCTAATGTG CAGGGATATT GCCTTATTTG   3480

TCTGTAAAAA ATGGAGCTCA GTAACATAAC TGCTTCTTGG AGCTTTGGAA TATTTTATCC   3540

TGTATTCTTG TTT   3553

(2) INFORMATION FOR SEQ ID NO:8:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 393 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: not relevant
        (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Ala Ala Pro Ser Pro Ser Gly Gly Gly Gly Ser Gly Thr Pro Gly
1   5   10   15

Pro Val Gly Ser Pro Ala Pro Gly His Pro Ala Val Ser Ser Met Gln
            20   25   30

```
Gly Lys Arg Lys Ala Leu Lys Leu Asn Phe Ala Asn Pro Pro Phe Lys
        35              40              45

Ser Thr Ala Arg Phe Thr Leu Asn Pro Asn Pro Thr Gly Val Gln Asn
    50              55              60

Pro His Ile Glu Arg Leu Arg Thr His Ser Ile Glu Ser Ser Gly Lys
65              70              75                      80

Leu Lys Ile Ser Pro Glu Gln His Trp Asp Phe Thr Ala Glu Asp Leu
            85              90              95

Lys Asp Leu Gly Glu Ile Gly Arg Gly Ala Tyr Gly Ser Val Asn Lys
            100             105             110

Met Val His Lys Pro Ser Gly Gln Ile Met Ala Val Lys Arg Ile Arg
        115             120             125

Ser Thr Val Asp Glu Lys Glu Gln Lys Gln Leu Leu Met Asp Leu Asp
    130             135             140

Val Val Met Arg Ser Ser Asp Cys Pro Tyr Ile Val Gln Phe Tyr Gly
145             150             155             160

Ala Leu Phe Arg Glu Gly Asp Cys Trp Ile Cys Met Glu Leu Met Ser
            165             170             175

Thr Ser Phe Asp Lys Phe Tyr Lys Tyr Val Tyr Ser Val Leu Asp Asp
            180             185             190

Val Ile Pro Glu Glu Ile Leu Gly Lys Ile Thr Leu Ala Thr Val Lys
            195             200             205

Ala Leu Met His Leu Lys Glu Asn Leu Lys Ile Ile His Arg Asp Ile
    210             215             220

Lys Pro Ser Asn Ile Leu Leu Asp Arg Ser Gly Met Ile Lys Leu Cys
225             230             235             240

Asp Phe Gly Ile Ser Gly Gln Leu Val Asp Ser Ile Ala Lys Thr Arg
            245             250             255

Asp Ala Gly Cys Arg Pro Tyr Met Ala Pro Glu Arg Ile Asp Phe Ser
            260             265             270

Ala Ser Arg Gln Gly Tyr Asp Val Arg Ser Asp Val Trp Ser Leu Gly
    275             280             285

Ile Thr Leu Tyr Glu Leu Ala Thr Gly Arg Phe Pro Tyr Pro Lys Trp
290             295             300

Asn Ser Val Phe Asp Gln Leu Thr Gln Val Val Lys Gly Asp Pro Pro
305             310             315             320

Gln Leu Ser Asn Ser Glu Glu Arg Glu Phe Ser Pro Ser Phe Ile Asn
            325             330             335

Phe Val Asn Leu Cys Leu Thr Lys Asp Glu Ser Lys Arg Pro Lys Tyr
            340             345             350

Lys Glu Leu Leu Lys His Pro Phe Ile Leu Met Tyr Glu Glu Arg Ala
    355             360             365
```

```
        Val Glu Val Ala Cys Tyr Val Cys Lys Ile Leu Asp Gln Met Pro Ala
        370                 375                 380

        Thr Pro Ser Ser Pro Met Tyr Val Asp
        385                 390
```

(2) INFORMATION FOR SEQ ID NO:9:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 3576 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
CTCCCAACAA TGGCGGCTCC GAGCCCGAGC GGCGGCGGCG GCTCCGGGGG CGGCAGCGGC    60

AGCGGCACCC CCGGCCCCGT AGGGTCCCCG GCGCCAGGCC ACCCGGCCGT CAGCAGCATG    120

CAGGGTAAAC GCAAAGCACT GAAGTTGAAT TTTGCAAATC CACCTTTCAA ATCTACAGCA    180

AGGTTTACTC TGAATCCCAA TCCTACAGGA GTTCAAAACC CACACATAGA GAGACTGAGA    240

ACACACAGCA TTGAGTCATC AGGAAAACTG AAGATCTCCC CTGAACAACA CTGGGATTTC    300

ACTGCAGAGG ACTTGAAAGA CCTTGGAGAA ATTGGACGAG GAGCTTATGG TTCTGTCAAC    360

AAAATGGTCC ACAAACCAAG TGGGCAAATA ATGGCAGTTA AAAGAATTCG GTCAACAGTG    420

GATGAAAAAG AACAAAAACA ACTTCTTATG GATTTGGATG TAGTAATGCG GAGTAGTGAT    480

TGCCCATACA TTGTTCAGTT TTATGGTGCA CTCTTCAGAG AGGGTGACTG TTGGATCTGT    540

ATGGAACTCA TGTCTACCTC GTTTGATAAG TTTTACAAAT ATGTATATAG TGTATTAGAT    600

GATGTTATTC CAGAAGAAAT TTTAGGCAAA ATCACTTTAG CAACTGTGAA AGCACTAAAC    660

CACTTAAAAG AAAACTTGAA AATTATTCAC AGAGATATCA AACCTTCCAA TATTCTTCTG    720

GACAGAAGTG GAAATATTAA GCTCTGTGAC TTCGGCATCA GTGGACAGCT TGTGGACTCT    780

ATTGCCAAGA CAAGAGATGC TGGCTGTAGG CCATACATGG CACCTGAAAG AATAGACCCA    840

AGCGCATCAC GACAAGGATA TGATGTCCGC TCTGATGTCT GGAGTTTGGG GATCACATTG    900

TATGAGTTGG CCACAGGCCG ATTTCCTTAT CCAAAGTGGA ATAGTGTATT TGATCAACTA    960

ACACAAGTCG TGAAAGGAGA TCCTCCGCAG CTGAGTAATT CTGAGGAAAG GGAATTCTCC    1020

CCGAGTTTCA TCAACTTTGT CAACTTGTGC CTTACGAAGG ATGAATCCAA AAGGCCAAAG    1080

TATAAAGAGC TTCTGAAACA TCCCTTTATT TTGATGTATG AAGAACGTGC CGTTGAGGTC    1140

GCATGCTATG TTTGTAAAAT CCTGGATCAA ATGCCAGCTA CTCCCAGCTC TCCCATGTAT    1200

GTCGATTGAT ATCGCTGCTA CATCAGACTC TAGAAAAAAG GGCTGAGAGG AAGCAAGACG    1260

TAAAGAATTT TCATCCCGTA TCACAGTGTT TTTATTGCTC GCCCAGACAC CATGTGCAAT    1320

AAGATTGGTG TTCGTTTCCA TCATGTCTGT ATACTCCTGT CACCTAGAAC GTGCATCCTT    1380
```

- 63 -

```
GTAATACCTG ATTGATCACA CAGTGTTAGT GCTGGTCAGA GAGACCTCAT CCTGCTCTTT      1440

TGTGATGAAC ATATTCATGA AATGTGGAAG TCAGTACGAT CAAGTTGTTG ACTGTGATTA      1500

GATCACATCT TAAATTCATT TCTAGACTCA AAACCTGGAG ATGCAGCTAC TGGAATGGTG      1560

TTTTGTCAGA CTTCCAAATC CTGGAAGGAC ACAGTGATGA ATGTACTATA TCTGAACATA      1620

GAAACTCGGG CTTGAGTGAG AAGAGCTTGC ACAGCCAACG AGACACATTG CCTTCTGGAG      1680

CTGGGAGACA AAGGAGGAAT TTACTTTCTT CACCAAGTGC AATAGATTAC TGATGTGATA      1740

TTCTGTTGCT TTACAGTTAC AGTTGATGTT TGGGGATCGA TGTGCTCAGC CAAATTTCCT      1800

GTTTGAAATA TCATGTTAAA TTAGAATGAA TTTATCTTTA CCAAAAACCA TGTTGCGTTC      1860

AAAGAGGTGA ACATTAAAAT ATAGAGACAG GACAGAATGT GTTCTTTTCT CCTCTACCAG      1920

TCCTATTTTT CAATGGGAAG ACTCAGGAGT CTGCCACTTG TCAAAGAAGG TGCTGATCCT      1980

AAGAATTTTT CATTCTCAGA ATTCGGTGTG CTGCCAACTT GATGTTCCAC CTGCCACAAA      2040

CCACCAGGAC TGAAAGAAGA AAACAGTACA GAAGGCAAAG TTTACAGATG TTTTTAATTC      2100

TAGTATTTTA TCTGGAACAA CTTGTAGCAG CTATATATTT CCCCTTGGTC CCAAGCCTGA      2160

TACTTTAGCC ATCATAACTC ACTAACAGGG AGAAGTAGCT AGTAGCAATG TGCCTTGATT      2220

GATTAGATAA AGATTTCTAG TAGGCAGCAA AAGACCAAAT CTCAGTTGTT TGCTTCTTGC      2280

CATCACTGGT CCAGGTCTTC AGTTTCCGAA TCTCTTTCCC TTCCCCTGTG GTCTATTGTC      2340

GCTATGTGAC TTGCGCTTAA TCCAATATTT TGCCTTTTTT CTATATCAAA AAACCTTTAC      2400

AGTTAGCAGG GATGTTCCTT ACCGAGGATT TTTAACCCCC AATCTCTCAT AATCGCTAGT      2460

GTTTAAAAGG CTAAGAATAG TGGGGCCCAA CCGATGTGGT AGGTGATAAA GAGGCATCTT      2520

TTCTAGAGAC ACATTGGACC AGATGAGGAT CCGAAACGGC AGCCTTTACG TTCATCACCT      2580

GCTAGAACCT CTCGTAGTCC ATCACCATTT CTTGGCATTG GAATTCTACT GGAAAAAAAT      2640

ACAAAAAGCA AAACAAAACC CTCAGCACTG TTACAAGAGG CCATTTAAGT ATCTTGTGCT      2700

TCTTCACTTA CCCATTAGCC AGGTTCTCAT TAGGTTTTGC TTGGGCCTCC CTGGCACTGA      2760

ACCTTAGGCT TTGTATGACA GTGAAGCAGC ACTGTGAGTG GTTCAAGCAC ACTGGAATAT      2820

AAAACAGTCA TGGCCTGAGA TGCAGGTGAT GCCATTACAG AACCAAATCG TGGCACGTAT      2880

TGCTGTGTCT CCTCTCAGAG TGACAGTCAT AAATACTGTC AAACAATAAA GGGAGAATGG      2940

TGCTGTTTAA AGTCACATCC CTGTAAATTG CAGAATTCAA AAGTGATTAT CTCTTTGATC      3000

TACTTGCCTC ATTTCCCTAT CTTCTCCCCC ACGGTATCCT AAACTTTAGA CTTCCCACTG      3060

TTCTGAAAGG AGACATTGCT CTATGTCTGC CTTCGACCAC AGCAAGCCAT CATCCTCCAT      3120

TGCTCCCGGG GACTCAAGAG GAATCTGTTT CTCTGCTGTC AACTTCCCAT CTGGCTCAGC      3180

ATAGGGTCAC TTTGCCATTA TGCAAATGGA GATAAAAGCA ATTCTGGCTG TCCAGGAGCT      3240
```

AATCTGACCG TTCTATTGTG TGGATGACCA CATAAGAAGG CAATTTTAGT GTATTAATCA   3300

TAGATTATTA TAAACTATAA ACTTAAGGGC AAGGAGTTTA TTACAATGTA TCTTTATTAA   3360

AACAAAAGGG TGTATAGTGT TCACAAACTG TGAAAATAGT GTAAGAACTG TACATTGTGA   3420

GCTCTGGTTA TTTTTCTCTT GTACCATAGA AAAATGTATA AAAATTATCA AAAAGCTAAT   3480

GTGCAGGGAT ATTGCCTTAT TTGTCTGTAA AAAATGGAGC TCAGTAACAT AACTGCTTCT   3540

TGGAGCTTTG GAATATTTTA TCCTGTATTC TTGTTT   3576

(2) INFORMATION FOR SEQ ID NO:10:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 399 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: not relevant
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Met Ala Ala Pro Ser Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Ser
1               5               10              15

Gly Ser Gly Thr Pro Gly Pro Val Gly Ser Pro Ala Pro Gly His Pro
            20                  25              30

Ala Val Ser Ser Met Gln Gly Lys Arg Lys Ala Leu Lys Leu Asn Phe
        35                  40              45

Ala Asn Pro Pro Phe Lys Ser Thr Ala Arg Phe Thr Leu Asn Pro Asn
        50                  55              60

Pro Thr Gly Val Gln Asn Pro His Ile Glu Arg Leu Arg Thr His Ser
65                  70              75              80

Ile Glu Ser Ser Gly Lys Leu Lys Ile Ser Pro Glu Gln His Trp Asp
                85                  90              95

Phe Thr Ala Glu Asp Leu Lys Asp Leu Gly Glu Ile Gly Arg Gly Ala
            100                 105             110

Tyr Gly Ser Val Asn Lys Met Val His Lys Pro Ser Gly Gln Ile Met
        115                 120             125

Ala Val Lys Arg Ile Arg Ser Thr Val Asp Glu Lys Glu Gln Lys Gln
        130                 135             140

Leu Leu Met Asp Leu Asp Val Val Met Arg Ser Ser Asp Cys Pro Tyr
145                 150             155             160

Ile Val Gln Phe Tyr Gly Ala Leu Phe Arg Glu Gly Asp Cys Trp Ile
                165             170             175

Cys Met Glu Leu Met Ser Thr Ser Phe Asp Lys Phe Tyr Lys Tyr Val
            180                 185             190

Tyr Ser Val Leu Asp Asp Val Ile Pro Glu Glu Ile Leu Gly Lys Ile
        195                 200             205

Thr Leu Ala Thr Val Lys Ala Leu Asn His Leu Lys Glu Asn Leu Lys
```

- 65 -

Ile Ile His Arg Asp Ile Lys Pro Ser Asn Ile Leu Leu Asp Arg Ser
225   230   235   240

Gly Asn Ile Lys Leu Cys Asp Phe Gly Ile Ser Gly Gln Leu Val Asp
                    245   250   255

Ser Ile Ala Lys Thr Arg Asp Ala Gly Cys Arg Pro Tyr Met Ala Pro
                260   265   270

Glu Arg Ile Asp Pro Ser Ala Ser Arg Gln Gly Tyr Asp Val Arg Ser
            275   280   285

Asp Val Trp Ser Leu Gly Ile Thr Leu Tyr Glu Leu Ala Thr Gly Arg
        290   295   300

Phe Pro Tyr Pro Lys Trp Asn Ser Val Phe Asp Gln Leu Thr Gln Val
305   310   315   320

Val Lys Gly Asp Pro Pro Gln Leu Ser Asn Ser Glu Glu Arg Glu Phe
                325   330   335

Ser Pro Ser Phe Ile Asn Phe Val Asn Leu Cys Leu Thr Lys Asp Glu
                340   345   350

Ser Lys Arg.Pro Lys Tyr Lys Glu Leu Leu Lys His Pro Phe Ile Leu
            355   360   365

Met Tyr Glu Glu Arg Ala Val Glu Val Ala Cys Tyr Val Cys Lys Ile
    370   375   380

Leu Asp Gln Met Pro Ala Thr Pro Ser Ser Pro Met Tyr Val Asp
385   390   395

(2) INFORMATION FOR SEQ ID NO:11:

    (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 393 amino acids
         (B) TYPE: amino acid
         (C) STRANDEDNESS: not relevant
         (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Pro Lys Lys Lys Pro Thr Pro Ile Gln Leu Asn Pro Ala Pro Asp
1   5   10   15

Gly Ser Ala Val Asn Gly Thr Ser Ser Ala Glu Thr Asn Leu Glu Ala
                20   25   30

Leu Gln Lys Lys Leu Glu Glu Leu Glu Leu Asp Glu Gln Gln Arg Lys
            35   40   45

Arg Leu Glu Ala Phe Leu Thr Gln Lys Gln Lys Val Gly Glu Leu Lys
        50   55   60

Asp Asp Asp Phe Glu Lys Ile Ser Glu Leu Gly Ala Gly Asn Gly Gly
65   70   75   80

Val Val Phe Lys Val Ser His Lys Pro Ser Gly Leu Val Met Ala Arg
                85   90   95

Lys Leu Ile His Leu Glu Ile Lys Pro Ala Ile Arg Asn Gln Ile Ile
                100   105   110

Arg Glu Leu Gln Val Leu His Glu Cys Asn Ser Pro Tyr Ile Val Gly
            115   120   125

Phe Tyr Gly Ala Phe Tyr Ser Asp Gly Glu Ile Ser Ile Cys Met Glu
    130   135   140

His Met Asp Gly Gly Ser Leu Asp Gln Val Leu Lys Lys Ala Gly Arg
145   150   155   160

Ile Pro Glu Gln Ile Leu Gly Lys Val Ser Ile Ala Val Ile Lys Gly
                165   170   175

Leu Thr Tyr Leu Arg Glu Lys His Lys Ile Met His Arg Asp Val Lys
            180   185   190

Pro Ser Asn Ile Leu Val Asn Ser Arg Gly Glu Ile Lys Leu Cys Asp
            195   200   205

Phe Gly Val Ser Gly Gln Leu Ile Asp Ser Met Ala Asn Ser Phe Val
    210   215   220

Gly Thr Arg Ser Tyr Met Ser Pro Glu Arg Leu Gln Gly Thr His Tyr
225   230   235   240

Ser Val Gln Ser Asp Ile Trp Ser Met Gly Leu Ser Leu Val Glu Met
            245   250   255

Ala Val Gly Arg Tyr Pro Ile Pro Pro Pro Asp Ala Lys Glu Leu Glu
            260   265   270

Leu Met Phe Gly Cys Gln Val Glu Gly Asp Ala Ala Glu Thr Pro Pro
    275   280   285

Arg Pro Arg Thr Pro Gly Arg Pro Leu Ser Ser Tyr Gly Met Asp Ser
    290   295   300

Arg Pro Pro Met Ala Ile Phe Glu Leu Leu Asp Tyr Ile Val Asn Glu
305   310   315   320

Pro Pro Pro Lys Leu Pro Ser Gly Val Phe Ser Leu Glu Phe Gln Asp
            325   330   335

Phe Val Asn Lys Cys Leu Ile Lys Asn Pro Ala Glu Arg Ala Asp Leu
            340   345   350

Lys Gln Leu Met Val His Ala Phe Ile Lys Arg Ser Asp Ala Glu Glu
            355   360   365

Val Asp Phe Ala Gly Trp Leu Cys Ser Thr Ile Gly Leu Asn Gln Pro
    370   375   380

Ser Thr Pro Thr His Ala Ala Gly Val
385   390

(2) INFORMATION FOR SEQ ID NO:12:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 400 amino acids
        (B) TYPE: amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
Met Leu Ala Arg Arg Lys Pro Val Leu Pro Ala Leu Thr Ile Asn Pro
1        .       5           10              15

Thr Ile Ala Glu Gly Pro Ser Pro Thr Ser Glu Gly Ala Ser Glu Ala
            20              25              30

Asn Leu Val Asp Leu Gln Lys Lys Leu Glu Glu Leu Glu Leu Asp Glu
        35              40              45

Gln Gln Lys Lys Arg Leu Glu Ala Phe Leu Thr Gln Lys Ala Lys Val
    50              55              60

Ser Glu Leu Lys Asp Asp Asp Phe Glu Arg Ile Ser Glu Leu Gly Ala
65              70              75              80

Gly Asn Gly Gly Val Val Thr Lys Val Gln His Arg Pro Ser Gly Leu
            85              90              95

Ile Met Ala Arg Lys Leu Ile His Leu Glu Ile Lys Pro Ala Ile Arg
        100             105             110

Asn Gln Ile Ile Arg Glu Leu Gln Val Leu His Glu Cys Asn Ser Pro
        115             120             125

Tyr Ile Val Gly Phe Tyr Gly Ala Phe Tyr Ser Asp Gly Glu Ile Ser
    130             135             140

Ile Cys Met Glu His Met Asp Gly Gly Ser Leu Asp Gln Val Leu Lys
145             150             155             160

Glu Ala Lys Arg Ile Pro Glu Glu Ile Leu Gly Lys Val Ser Ile Ala
            165             170             175

Val Leu Arg Gly Leu Ala Tyr Leu Arg Glu Lys His Gln Ile Met His
            180             185             190

Arg Asp Val Lys Pro Ser Asn Ile Leu Val Asn Ser Arg Gly Glu Ile
        195             200             205

Lys Leu Cys Asp Phe Gly Val Ser Gly Gln Leu Ile Asp Ser Met Ala
    210             215             220

Asn Ser Phe Val Gly Thr Arg Ser Tyr Met Ala Pro Glu Arg Leu Gln
225             230             235             240

Gly Thr His Tyr Ser Val Gln Ser Asp Ile Trp Ser Met Gly Leu Ser
            245             250             255

Leu Val Glu Leu Ala Val Gly Arg Tyr Pro Ile Pro Pro Asp Ala
        260             265             270

Lys Glu Leu Glu Ala Ile Phe Gly Arg Pro Val Val Asp Gly Glu Glu
    275             280             285

Gly Glu Pro His Ser Ile Ser Pro Arg Pro Arg Pro Pro Gly Arg Pro
    290             295             300
```

- 68 -

```
Val Ser Gly His Gly Met Asp Ser Arg Pro Ala Met Ala Ile Phe Glu
305             310             315             320

Leu Leu Asp Tyr Ile Val Asn Glu Pro Pro Pro Lys Leu Pro Asn Gly
            325             330             335

Val Phe Thr Pro Asp Phe Gln Glu Phe Val Asn Lys Cys Leu Ile Lys
        340             345             350

Asn Pro Ala Glu Arg Ala Asp Leu Lys Met Leu Thr Asn His Thr Phe
        355             360             365

Ile Lys Arg Ser Glu Val Glu Glu Val Asp Phe Ala Gly Trp Leu Cys
    370             375             380

Lys Thr Leu Arg Leu Asn Gln Pro Gly Thr Pro Thr Arg Thr Ala Val
385             390             395             400
```

(2) INFORMATION FOR SEQ ID NO:13:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 668 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: not relevant
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Met Glu Asp Lys Phe Ala Asn Leu Ser Leu His Glu Lys Thr Gly Lys
1               5               10              15

Ser Ser Ile Gln Leu Asn Glu Gln Thr Gly Ser Asp Asn Gly Ser Ala
        20              25              30

Val Lys Arg Thr Ser Ser Thr Ser Ser His Tyr Asn Asn Ile Asn Ala
        35              40              45

Asp Leu His Ala Arg Val Lys Ala Phe Gln Glu Gln Arg Ala Leu Lys
    50              55              60

Arg Ser Ala Ser Val Gly Ser Asn Gln Ser Glu Gln Asp Lys Gly Ser
65              70              75              80

Ser Gln Ser Pro Lys His Ile Gln Gln Ile Val Asn Lys Pro Leu Pro
            85              90              95

Pro Leu Pro Val Ala Gly Ser Ser Lys Val Ser Gln Arg Met Ser Ser
        100             105             110

Gln Val Val Gln Ala Ser Ser Lys Ser Thr Leu Lys Asn Val Leu Asp
        115             120             125

Asn Gln Glu Thr Gln Asn Ile Thr Asp Val Asn Ile Asn Ile Asp Thr
    130             135             140

Thr Lys Ile Thr Ala Thr Thr Ile Gly Val Asn Ile Gly Leu Pro Ala
145             150             155             160

Thr Asp Ile Thr Pro Ser Val Ser Asn Thr Ala Ser Ala Thr His Lys
            165             170             175
```

- 69 -

```
Ala Gln Leu Leu Asn Pro Asn Arg Arg Ala Pro Arg Arg Pro Leu Ser
            180             185             190

Thr Gln His Pro Thr Arg Pro Asn Val Ala Pro His Lys Ala Pro Ala
        195             200             205

Ile Ile Asn Thr Pro Lys Gln Ser Leu Ser Ala Arg Arg Gly Leu Lys
    210             215             220

Leu Pro Pro Gly Gly Met Ser Leu Lys Met Pro Thr Lys Thr Ala Gln
225             230             235             240

Gln Pro Gln Gln Phe Ala Pro Ser Pro Ser Asn Lys Lys His Ile Glu
            245             250             255

Thr Leu Ser Asn Ser Lys Val Val Glu Gly Lys Arg Ser Asn Pro Gly
        260             265             270

Ser Leu Ile Asn Gly Val Gln Ser Thr Ser Thr Ser Ser Ser Thr Glu
    275             280             285

Gly Pro His Asp Thr Val Gly Thr Thr Pro Arg Thr Gly Asn Ser Asn
    290             295             300

Asn Ser Ser Asn Ser Gly Ser Ser Gly Gly Gly Gly Leu Phe Ala Asn
305             310             315             320

Phe Ser Lys Tyr Val Asp Ile Lys Ser Gly Ser Leu Asn Phe Ala Gly
            325             330             335

Lys Leu Ser Leu Ser Ser Lys Gly Ile Asp Phe Ser Asn Gly Ser Ser
        340             345             350

Ser Arg Ile Thr Leu Asp Glu Leu Glu Phe Leu Asp Glu Leu Gly His
        355             360             365

Gly Asn Tyr Gly Asn Val Ser Lys Val Leu His Lys Pro Thr Asn Val
    370             375             380

Ile Met Ala Thr Lys Glu Val Arg Leu Glu Leu Asp Glu Ala Lys Phe
385             390             395             400

Arg Gln Ile Leu Met Glu Leu Glu Val Leu His Lys Cys Asn Ser Pro
            405             410             415

Tyr Ile Val Asp Phe Tyr Gly Ala Phe Phe Ile Glu Gly Ala Val Tyr
            420             425             430

Met Cys Met Glu Tyr Met Asp Gly Gly Ser Leu Asp Lys Ile Tyr Asp
        435             440             445

Glu Ser Ser Glu Ile Gly Gly Ile Asp Glu Pro Gln Leu Ala Phe Ile
    450             455             460

Ala Asn Ala Val Ile His Gly Leu Lys Glu Leu Lys Glu Gln His Asn
465             470             475             480

Ile Ile His Arg Asp Val Lys Pro Thr Asn Ile Leu Cys Ser Ala Asn
            485             490             495

Gln Gly Thr Val Lys Leu Cys Asp Phe Gly Val Ser Gly Asn Leu Val
            500             505             510
```

- 70 -

```
Ala Ser Leu Ala Lys Thr Asn Ile Gly Cys Gln Ser Tyr Met Ala Pro
        515             520             525

Glu Arg Ile Lys Ser Leu Asn Pro Asp Arg Ala Thr Tyr Thr Val Gln
    530             535             540

Ser Asp Ile Trp Ser Leu Gly Leu Ser Ile Leu Glu Met Ala Leu Gly
545             550             555             560

Arg Tyr Pro Tyr Pro Pro Glu Thr Tyr Asp Asn Ile Phe Ser Gln Leu
            565             570             575

Ser Ala Ile Val Asp Gly Pro Pro Pro Arg Leu Pro Ser Asp Lys Phe
        580             585             590

Ser Ser Asp Ala Gln Asp Phe Val Ser Leu Cys Leu Gln Lys Ile Pro
    595             600             605

Glu Arg Arg Pro Thr Tyr Ala Ala Leu Thr Glu His Pro Trp Leu Val
    610             615             620

Lys Tyr Arg Asn Gln Asp Val His Met Ser Glu Tyr Ile Thr Glu Arg
625             630             635             640

Leu Glu Arg Arg Asn Lys Ile Leu Arg Glu Arg Gly Glu Asn Gly Leu
            645             650             655

Ser Lys Asn Val Pro Ala Leu His Met Gly Gly Leu
        660             665
```

(2) INFORMATION FOR SEQ ID NO:14:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 23 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTYTAYGGNG CNTTYTTYAT HGA   23

(2) INFORMATION FOR SEQ ID NO:15:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 20 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATBCTYTCNG GNGCCATKTA   20

(2) INFORMATION FOR SEQ ID NO:16:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 17 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ASTYRYSASA SASASYS